

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-344567

(43)Date of publication of application : 14.12.2001

(51)Int.Cl. ....

(21)Application number : 2000-166691 (71)Applicant : RICOH CO LTD

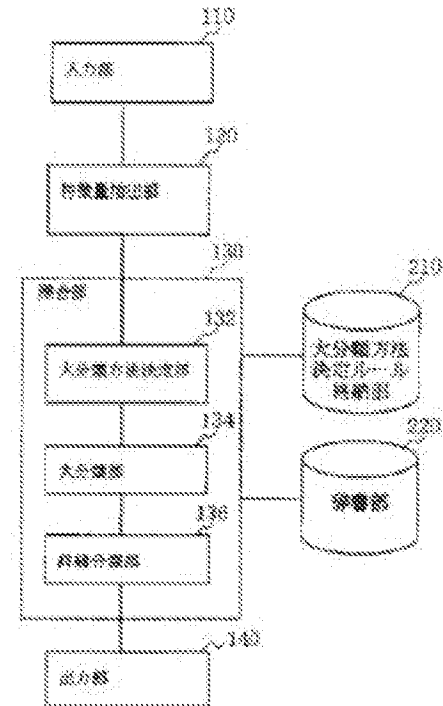
(22)Date of filing : 02.06.2000 (72)Inventor : SEN RITSU

## (54) DEVICE AND METHOD FOR RECOGNIZING CHARACTER AND RECORDING MEDIUM WITH PROGRAM FOR PERFORMING THE METHOD RECORDED THEREON

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To provide a character recognizing device which fast and accurately performs character recognition processing by adopting a roughly classifying method that is suitable to the types of recognition dictionary parts and the image attribute of a collation object character in the case of using various recognition dictionary parts.

**SOLUTION:** This device is provided with an inputting part 110 for inputting a character image, a characteristic quantity extracting part 120 for extracting characteristic quantity from the character image inputted by the part 110, a roughly classifying method deciding part 132 for deciding a roughly classifying method according to information for the inputted character image, a roughly classifying part 134 for narrowing down candidate characters with the roughly classifying method decided by the part 132 and the characteristic quantity extracted by the part 120, a detailed classifying part 136 for calculating the degree of difference of the candidate character narrowed down by the part 134 from a dictionary part 220 for collation and an outputting part 140 for outputting a character category obtained by the part 136 as a candidate.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号  
特開2001-344567  
(P2001-344567A)

(43) 公開日 平成13年12月14日 (2001. 12. 14)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	データ形式 (参考)
G 0 6 K 9/62	6 2 0	C 0 6 K 9/62	6 2 0 A 5 B 0 6 4
	6 1 0		6 1 0 C
			6 1 0 D
9/68		9/68	B

審査請求 未請求 請求項の数10 O L (全 8 頁)

(21) 出願番号 特願2000-166691(P2000-166691)

(22) 出願日 平成12年6月2日 (2000. 6. 2)

(71) 出願人 000006747

株式会社リコー

東京都大田区中馬込1丁目3番6号

(72) 発明者 ▲せん▼ 立

東京都大田区中馬込1丁目3番6号 株式会社リコー内

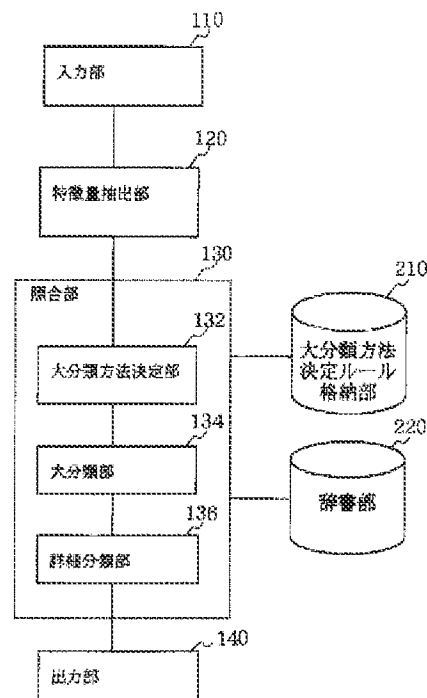
Fターム(参考) 5B064 AB13 AB14 AB16 AB18 DA06  
DA09 DA14 DA16 DA35

(54) 【発明の名称】 文字認識装置、その方法およびその方法を実施するためのプログラムを記録した記録媒体

(57) 【要約】

【課題】 種々の認識辞書を使う場合、この認識辞書の種類や照合対象文字の画像属性に適應した大分類方法をとることによって、高速かつ高精度に文字の認識処理を実施する文字認識装置およびその方法を提供する。

【解決手段】 本発明は、文字画像を入力する入力部110と、この入力部110で入力された文字画像から特徴量を抽出する特徴量抽出部120と、入力された文字画像に対する情報によって大分類方法を決定する大分類方法決定部132と、この大分類方法決定部132で決定された大分類方法と特徴量抽出部120で抽出された特徴量とで候補文字を絞る大分類部134と、この大分類部134で絞り込まれた候補文字に対して照合用の辞書部220との相違度を計算する詳細分類部136と、この詳細分類部136で得た文字カテゴリを候補として出力する出力部140とを備える。



## 【特許請求の範囲】

【請求項1】 入力した文字画像を大分類で候補文字を絞り、その候補文字について辞書部と照合して照合結果を出力する文字認識装置において、前記候補文字を絞る際、入力された文字画像に対する情報によって大分類方法を決定する大分類方法決定部を備えるようにしたことを特徴とする文字認識装置。

【請求項2】 文字画像を入力する入力部と、前記入力部で入力された文字画像から特徴量を抽出する特徴量抽出部と、入力された文字画像に対する情報によって大分類方法を決定する大分類方法決定部と、前記大分類方法決定部で決定された大分類方法と前記特徴量抽出部で抽出された特徴量とで候補文字を絞る大分類部と、前記大分類部で絞り込まれた候補文字に対して照合用の辞書部と相違度を計算する詳細分類部と、前記詳細分類部で得た文字カテゴリを候補として出力する出力部とを備えたことを特徴とする文字認識装置。

【請求項3】 請求項1または請求項2記載の文字認識装置において、前記大分類方法決定部は、前記文字画像に対する情報として、指定された辞書部の種類または照合用の辞書部中の照合対象文字属性、あるいは、照合対象文字の画像属性を有し、これらの情報に対して大分類を行うか否かを決定し、大分類を行う場合には、その大分類方法を対応付けた大分類決定ルールを格納している大分類決定ルール格納部を備えるようにして、大分類方法を決定するようにしたことを特徴とする文字認識装置。

【請求項4】 請求項1、請求項2または請求項3記載の文字認識装置において、前記大分類方法は、照合の距離による大分類、候補順位による大分類、照合の距離と画像属性による大分類、または、これらの方法を組み合わせた大分類からなることを特徴とする文字認識装置。

【請求項5】 請求項3記載の文字認識装置において、前記辞書部の種類は、漢字だけの辞書部、仮名だけの辞書部、英字だけの辞書部、記号だけの辞書部、縦長文字だけの辞書部、横長文字の辞書部、小文字だけの辞書部、または、これらが混在した辞書部の1つ以上の辞書部を備えたことを特徴とする文字認識装置。

【請求項6】 文字画像を入力する入力工程と、前記入力工程で入力された文字画像から特徴量を抽出する特徴量抽出工程と、入力された文字画像に対する情報によって大分類方法を決定する大分類方法決定工程と、前記大分類方法決定工程で決定された大分類方法と前記特徴量抽出工程で抽出された特徴量とで候補文字を絞る大分類工程と、前記大分類部で絞り込まれた候補文字に対して照合用の辞書部と相違度を計算する詳細分類工程と、前記詳細分類工程で得た文字カテゴリを候補として出力する出力工程とを備えたことを特徴とする文字認識方法。

【請求項7】 請求項6記載の文字認識方法において、前記文字画像に対する情報として、指定された辞書部の

種類または照合用の辞書部中の照合対象文字属性、あるいは、照合対象文字の画像属性を有し、これらの情報に対して大分類を行うか、否かを決定し、大分類を行う場合には、その大分類方法を対応付けた大分類決定ルールを格納している大分類決定ルール格納部を備えるようにして、大分類方法を決定するようにしたことを特徴とする文字認識方法。

【請求項8】 請求項6または請求項7記載の文字認識方法において、前記大分類方法は、照合の距離による大分類、候補順位による大分類、照合の距離と画像属性による大分類、または、これらの方法を組み合わせた大分類からなることを特徴とする文字認識方法。

【請求項9】 請求項7記載の文字認識方法において、前記辞書部の種類は、漢字だけの辞書部、仮名だけの辞書部、英字だけの辞書部、記号だけの辞書部、縦長文字だけの辞書部、横長文字の辞書部、小文字だけの辞書部、または、これらが混在した辞書部の1つ以上の辞書部を備えたことを特徴とする文字認識方法。

【請求項10】 文字画像を入力する入力機能と、前記入力機能により入力された文字画像から特徴量を抽出する特徴量抽出機能と、入力された前記文字画像に対する情報によって大分類方法を決定する大分類方法決定機能と、前記大分類方法決定機能により決定された大分類方法と前記特徴量抽出機能により抽出された特徴量とで候補文字を絞る大分類機能と、前記大分類機能により絞り込まれた候補文字に対して照合用の辞書部との相違度を計算する詳細分類機能と、前記詳細分類機能で得た文字カテゴリを候補として出力する出力機能部とを実施するためのプログラムを記録したコンピュータ読み取り可能な記録媒体。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、文字を認識するための文字認識装置、その方法およびその方法を実施するためのプログラムを記録した記録媒体に関し、特に、文字を認識するとき大分類を行って候補を絞り込んでより詳細な認識を行うための文字認識装置、その方法およびその方法を実施するためのプログラムを記録した記録媒体に関する。

## 【0002】

【従来の技術】従来、文字認識装置においては、入力文字パターンとカテゴリごとに作成した辞書部との照合を行うことにより、認識する方法が多く採られている。この場合、認識対象となる字種が数字や記号等のように数十のオーダーである場合には入力文字パターンとすべてのカテゴリとの間で照合計算をしても比較的短い時間に結果を得られるが、漢字のように字種が数千のオーダーである場合には数字や記号等の計算時間に比較して百倍以上の時間がかかる。また、これにフォントや字体等の変化も考慮すると、この照合計算には膨大な時間がか

かることになる。そこで照合時間を短くするために、大分類を行って候補を比較的少数に絞り、次に絞られた候補に対して詳細な照合計算を行うという処理が採られている。この大分類を行う方法としては、例えば、大分類用辞書部と詳細分類用辞書部とを別々に備えて、文字画像より抽出した大分類用特徴量と大分類用辞書部との照合計算（大分類）を行うことにより、詳細分類用辞書部との照合計算（詳細分類）を行うか否か、あるいは、照合する範囲を決定している（特開昭63-126082号公報等）。

【0003】

【発明が解決しようとする課題】しかしながら、この大分類で候補文字数を絞り込んで行う従来の文字認識方法においては、ユーザの用途や文字認識のアルゴリズムによって、種々の認識辞書部を使う場合、この認識辞書部の種類にかかわらず、共通の方法で大分類するのでは、認識速度や認識精度の低下が懸念されるという問題がある。本発明の課題は、このような問題を解決することにある。すなわち、本発明の目的は、辞書部の種類や照合対象文字の画像属性に適應した分類方法をとることにより、文字認識を高速かつ高精度に実施することができる文字認識装置、その方法およびその方法を実施するためのプログラムを記録した記録媒体を提供することにある。

【0004】

【課題を解決するための手段】本発明の請求項1の文字認識装置は、入力した文字画像を大分類で候補文字を絞り、その候補文字について辞書部と照合して照合結果を出力する文字認識装置において、前記候補文字を絞る際、入力された文字画像に対する情報によって大分類方法を決定する大分類方法決定部を備えるようにしたことを特徴とする。また、本発明の請求項2の文字認識装置は、文字画像を入力する入力部と、前記入力部で入力された文字画像から特徴量を抽出する特徴量抽出部と、入力された文字画像に対する情報によって大分類方法を決定する大分類方法決定部と、前記大分類方法決定部で決定された大分類方法と前記特徴量抽出部で抽出された特徴量とで候補文字を絞る大分類部と、前記大分類部で絞り込まれた候補文字に対して照合用の辞書部と相違度を計算する詳細分類部と、前記詳細分類部で得た文字カテゴリを候補として出力する出力部とを備えたことを特徴とする。また、本発明の請求項3の文字認識装置は、請求項1または請求項2記載の文字認識装置において、前記大分類方法決定部は、前記文字画像に対する情報として、指定された辞書部の種類または照合用の辞書部中の照合対象文字属性、あるいは、照合対象文字の画像属性を有し、これらの情報に対して大分類を行うか、否かを決定し、大分類を行う場合には、その大分類方法に対応付けた大分類決定ルールを備えるようにして、大分類方法を決定するようにしたことを特徴とする。また、本発

明の請求項4の文字認識装置は、請求項1、請求項2、または請求項3記載の文字認識装置において、前記大分類方法は、照合の距離による大分類、候補順位による大分類、照合の距離と画像属性による大分類、または、これらの方法を組み合わせた大分類からなることを特徴とする。また、本発明の請求項5の文字認識装置は、請求項3記載の文字認識装置において、前記辞書部の種類は、漢字だけの辞書部、仮名だけの辞書部、英字だけの辞書部、記号だけの辞書部、縦長文字だけの辞書部、横長文字の辞書部、小文字だけの辞書部、または、これらが混在した辞書部の1つ以上の辞書部を備えたことを特徴とする。

【0005】また、本発明の請求項6の文字認識方法は、文字画像を入力する入力工程と、前記入力工程で入力された文字画像から特徴量を抽出する特徴量抽出工程と、入力された文字画像に対する情報によって大分類方法を決定する大分類方法決定工程と、前記大分類方法決定工程で決定された大分類方法と前記特徴量抽出工程で抽出された特徴量とで候補文字を絞る大分類工程と、前記大分類部で絞り込まれた候補文字に対して照合用の辞書部と相違度を計算する詳細分類工程と、前記詳細分類工程で得た文字カテゴリを候補として出力する出力工程とを備えたことを特徴とする。また、本発明の請求項7の文字認識方法は、請求項6記載の文字認識方法において、前記文字画像に対する情報として、指定された辞書部の種類または照合用の辞書部中の照合対象文字属性、あるいは、照合対象文字の画像属性を有し、これらの情報に対して大分類を行うか、否かを決定し、大分類を行う場合には、その大分類方法に対応付けた大分類決定ルールを備えるようにして、大分類方法を決定するようにしたことを特徴とする。また、本発明の請求項8の文字認識方法は、請求項6または請求項7記載の文字認識方法において、前記大分類方法は、照合の距離による大分類、候補順位による大分類、照合の距離と画像属性による大分類、または、これらの方法を組み合わせた大分類からなることを特徴とする。また、本発明の請求項9の文字認識方法は、請求項7記載の文字認識方法において、前記辞書部の種類は、漢字だけの辞書部、仮名だけの辞書部、英字だけの辞書部、記号だけの辞書部、縦長文字だけの辞書部、横長文字の辞書部、小文字だけの辞書部、または、これらが混在した辞書部の1つ以上の辞書部を備えたことを特徴とする。また、本発明の請求項10の記録媒体は、コンピュータを、文字認識装置として機能させるためのコンピュータ読み取り可能な記録媒体であって、このコンピュータを、文字画像を入力する入力部と、前記入力部で入力された文字画像から特徴量を抽出する特徴量抽出部と、入力された文字画像に対する情報によって大分類方法を決定する大分類方法決定部と、前記大分類方法決定部で決定された大分類方法と前記特徴量抽出部で抽出された特徴量とで候補文字を絞る

大分類部と、前記大分類部で絞り込まれた候補文字に対して照合用の辞書部と相違度を計算する詳細分類部と、前記詳細分類部で得た文字カテゴリを候補として出力する出力部とを備えた文字認識装置として機能させるためのプログラムを記録した。

【0006】

【発明の実施の形態】以下、本発明の1つの実施の形態を図面に基いて詳細に説明する。

(1) 文字認識装置の機能構成

図1は、本発明の1つの実施の形態に係る文字認識装置の機能構成を示すブロック図である。本発明の1つの実施の形態に係る文字認識装置は、入力部110、特徴量抽出部120、照合部130、出力部140及び照合用の辞書部220から構成されている。更に、照合部130は、大分類方法決定部132、大分類部134、詳細分類部136及び大分類方法決定ルール格納部210とから構成されている。

(2) 文字認識の概略的な処理の流れ

入力部110は、次のような文字認識に必要な前処理を行う。まず、原稿用紙等の画像をスキャナーによって読み取る。画像の読み取りは、スキャナーだけではなく、既に読み込まれている画像をファイル等から読み込んでもよい。この読み込まれた画像をディスプレイへ表示し、その中からマウスのようなポインティングデバイスから文字部分を選択する。文字部分は、領域識別の技術を使って判別されても良い。この選択された文字部分の画像に対して大きさの正規化や雑音除去を行う。これらの処理が終了すると、特徴量抽出部120へ送られる。この入力された文字の画像属性として、面積、形状(縦長、横長等)、配置(上付き、下付き、真中等)、画素の密度(白と黒の割合)等を抽出して、一時的に記憶しておく。また、文字の画像特性として、面積、形状(縦長、横長等)、配置(上付き、下付き、真中等)から自動的にどの辞書部と照合するかを判断する。

【0007】この照合用の辞書部220としては、図3に示したように漢字だけの辞書部、仮名だけの辞書部、英字だけの辞書部、記号だけの辞書部、縦長文字(例、“1”、“i”等)だけの辞書部、横長文字(例、“—”、“—”等)の辞書部、小文字(例：“。”、“、”、“・”等の面積的に小さな文字)だけの辞書部、または、それら混在している辞書部等が用意されており、この中から1乃至複数の辞書部を指定して照合する。特徴量抽出部120は、入力部110から送られた文字画像に対して、認識に用いられる特徴量である特徴ベクトルを抽出する。この特徴ベクトルは、特開平1-250184号公報の方向コードヒストグラム等を用いて作成される。照合部130は、入力部110で指示された辞書部の種類や照合対象文字の画像属性に対応する大分類方法を、大分類方法決定ルール格納部210(図5参照)の大分類方法決定ルールを用いて大

分類方法決定部132(後述)で決定し、その大分類方法によって入力文字を大分類部134で大雑把に分類分けして候補を絞り、この候補について詳細分類部136が、抽出された特徴ベクトルと照合用の辞書部220に登録された標準パターンの特徴ベクトルとを所定の計算式を用いて相違度を計算する。この相違度の計算は、例えば、ユークリッド距離や擬似ベイズ識別関数等によって計算される。また、この大分類方法には、図4にあるように、マッチングの距離による大分類方法、候補順位による大分類方法、マッチングの距離と画像属性による大分類方法、組み合わせによる大分類方法および大分類を行わない方法とがある。出力部140は、照合対象文字に対する照合結果である相違度を小さい順に並べ、そのうちの一定個数(例えば、10個)につき文字コードを出力する。照合用の辞書部220は、上記したような種々の種類があり、それらは登録されている文字毎に、文字コード、標準パターン特徴ベクトル等の情報を保持している。

【0008】(3) 本発明の1つの実施の形態における処理の流れ

図2は、本発明の1つの実施の形態における処理の流れを示している。原稿用紙等の画像をスキャナーによって読み取る。この画像の読み取りは、スキャナーだけではなく、既に読み込まれている画像をファイル等から読み込んでもよい。この読み込まれた画像をディスプレイへ表示し、その中からマウスのようなポインティングデバイスから文字部分を選択する。文字部分は、領域識別の技術を使って判別されても良い。この選択された文字部分の画像に対して大きさの正規化や雑音除去を行う。また、この入力された文字の画像属性として、面積、形状(縦長、横長等)、配置(上付き、下付き、真中等)、画素の密度(白と黒の割合)等を抽出して、一時的に記憶しておく(ステップS100)。この処理する文字をどの照合用の辞書部(図3参照)によって照合処理するかは、この文字の画像特性として、面積、形状(縦長、横長等)、配置(上付き、下付き、真中等)等から判断する。この選択結果を一時的に記憶しておく(ステップS110)。入力された文字画像に対して、認識に用いられる特徴量である特徴ベクトルを抽出する(ステップS120)。一時的に保存された照合に使う辞書部の種類や照合対象文字の画像属性に対応する大分類方法を、大分類方法決定ルール210(図5参照)に適用して大分類方法を決定する(ステップS130)。この決定された大分類方法(図4参照)によって、入力文字を大雑把に分類分けして候補を絞る(ステップS140)。この候補について、抽出された特徴ベクトルと指定された照合用の辞書部220に登録された標準パターンの特徴ベクトルとを所定の計算式を用いて相違度を計算して、最終的な候補の絞り込みを行う。この相違度の計算は、例えば、ユークリッド距離や擬似ベイズ識別関数等によ

て計算される(ステップS150)。照合対象文字に対する照合結果である相違度を小さい順に並べ、そのうちの一定個数につき文字コードを最終候補として出力する(ステップS160)。

【0009】(4)大分類方法決定部132の説明

(A)辞書部の種類を指定された場合

辞書部エントリ数の少ない小文字

(例、“。”、“、”、“、”等の面積の小さな文字)だけの辞書部を使った場合、大分類によって、候補数は0になり、正しい認識結果が得られない場合がある。このような場合には、辞書部の種類によって、大分類を行わないか、または大分類の方法を別のものに代えたほうが良い。例えば、図5に示すように、小文字だけの辞書部を指定された場合は、「大分類を使わない」というようにルールを定める。また、混在辞書部の場合は、「マッチング距離による大分類と候補順位による大分類の両方を用いる」ようにルールを定める。

(B)照合対象文字の画像属性の場合

入力された文字画像の画像属性として、一時的に記憶した面積、形状(縦長、横長等)、配置(上付き、下付き、真中等)、画素の密度(白と黒の割合)等によって、同じ辞書部の中で照合対象の辞書部エントリを指定することがある。例えば、入力画像の矩形面積が小さい場合は、辞書部中の小文字(“。”、“、”等)エントリだけを照合対象とする。または、この場合、小文字だけの辞書部を対象とする。また、画素の密度によっては漢字だけの辞書部を指定されたとみなしたり、形状によっては縦長文字だけの辞書部や横長文字だけの辞書部が指定されたものとみなすことができる。このように対象を限定されたときには、上記した辞書部の種類を指定されたときと同様に大分類方法決定ルール格納部210の大分類方法決定ルールを適用する。

【0010】次に、本発明の他の実施の形態を図面に基づいて説明する。図6は本発明の他の実施の形態に係る文字認識装置を示すブロック図である。本発明の他の実施の形態に係る文字認識装置は、図6に示したハードウェア構成を持つコンピュータ装置により構成されている。このコンピュータ装置は、入力装置1と、表示装置2と、CPU3と、メモリ4と、記憶装置5と、媒体駆動装置6とを有している。入力装置1と、表示装置2と、CPU3と、メモリ4と、記憶装置5と、媒体駆動装置6は、相互に接続されている。入力装置1はキーボード、マウス、タッチパネル等により構成され、情報の入力に使用される。また、文字画像を入力するためにスキャナが接続されても良い。表示装置2は、種々の出力情報や入力装置1からの入力された情報などを表示出力させる。CPU(Central Processing Unit; 中央処理ユニット)3は、種々のプログラムを動作させる。メモリ4は、プログラム自身を保持し、またそのプログラムがCPU3によって実行されるときに一時的に作成さ

れる情報等を保持する。記憶装置5は、本装置で扱う大分類方法決定ルール格納部210や照合用の辞書部220及びプログラムやプログラム実行時の一時的な情報等を保持する。媒体駆動装置6は、プログラムやデータ等を記憶した記録媒体を装着してそれらを読み込み、メモリ4または記憶装置5へ格納するのに用いられる。また、直接データ(例えば、文字認識対象の文書画像データ等)の入出力やプログラム実行するのに使ってもよい。このようなコンピュータ装置において、図1に示した文字認識装置を構成する各機能をプログラム化し、予めCD-ROM等の記録媒体に書き込んでおき、このCD-ROMを各サイトのCD-ROMドライブのような媒体駆動装置6を搭載したコンピュータに装着して、プログラムをメモリ4あるいは記憶装置5に格納し、そのプログラムを実行することによって、前記実施形態と同様な機能を実施することができる。

【0011】なお、記録媒体としては半導体媒体(例えば、ROM、ICメモリカード等)、光媒体(例えば、DVD-ROM、MO、MD、CD-R等)、磁気媒体(例えば、磁気テープ、フレキシブルディスク等)のいずれであってもよい。また、本発明の機能を実施するプログラムは、媒体の形で頒布することができる。また、本発明の機能を実施するプログラムを磁気ディスク等の記憶装置に格納しておき、有線または無線の通信ネットワークによりダウンロード等の形式で頒布することも可能である。更に、本発明の機能を実施するプログラムを放送波によって配布することによって提供するようにしても良い。また、クライアント側のコンピュータのスキャナから画像を取り込み、取り込んだ画像をサーバー側の入力部110へ送り、認識処理を行い、認識結果を出力部140がクライアント側のコンピュータへ送信して、クライアント側のコンピュータでその送られた認識結果を表示するようにさせても良い。また、クライアント側のコンピュータで取り込んだ画像の初期処理を行ってから、サーバー側の入力部110へその処理結果を送るようにしても良い。

【0012】

【発明の効果】以上のように、本発明によれば、照合対象文字に関する情報(指定された辞書部の種類や照合対象文字の画像属性)によって、適切な大分類方法へ切り替えることができるので、高速かつ高精度に文字の認識処理を実施することができる。

【図面の簡単な説明】

【図1】本発明の1つの実施の形態に係る文字認識装置を示すブロック図である。

【図2】図1の文字認識装置の動作を説明するためのフローチャートである。

【図3】本発明の1つの実施の形態で用いる辞書部の種類を説明するための図である。

【図4】本発明の1つの実施の形態で用いる大分類方法

の種類を説明するための図である。

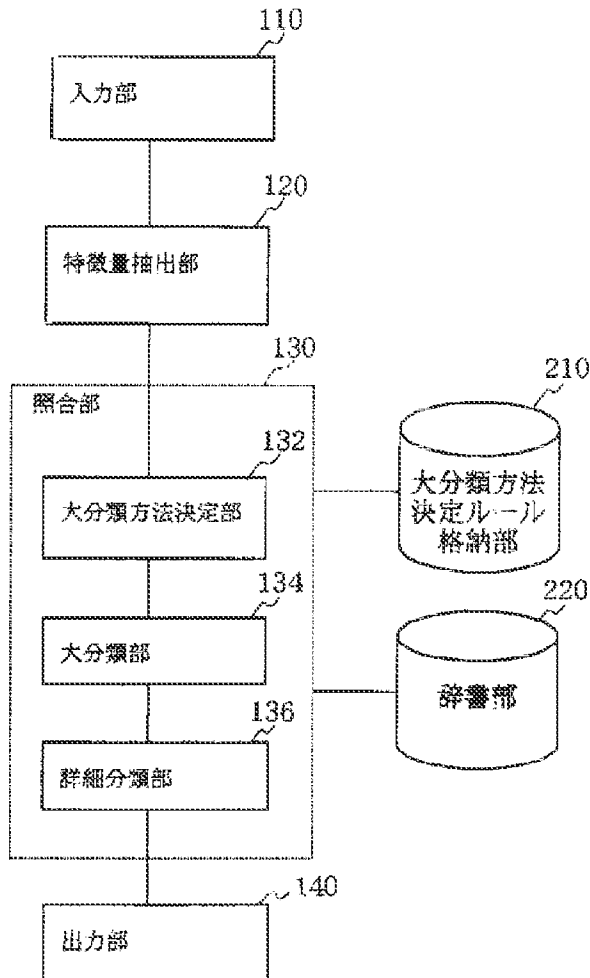
【図5】本発明の1つの実施の形態で用いる大分類を決定するためのルールを説明する図である。

【図6】本発明の他の実施の形態に係る文字認識装置を示すブロック図である。

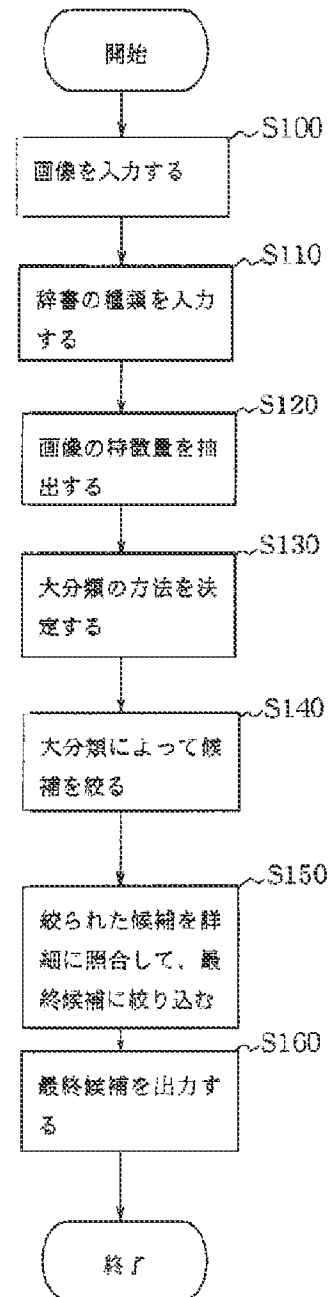
【符号の説明】

110 入力部、120 特徴量抽出部、130 照合部、132 大分類方法決定部、134 大分類部、136 詳細分類部、140 出力部、210 大分類方法決定ルール格納部、220 辞書部、1 入力装置、2 表示装置、3 CPU、4 メモリ、5 記憶装置、6 媒体駆動装置。

【図1】



【図2】



【図3】

## 辞書の種類

辞書1	漢字だけの辞書
辞書2	仮名だけの辞書
辞書3	英字だけの辞書
辞書4	数字だけの辞書
辞書5	記号だけの辞書
辞書6	縦長文字だけの辞書
辞書7	横長文字だけの辞書
辞書8	小文字だけの辞書
辞書9	漢字、仮名、英字、記号などの混在辞書
辞書10	辞書の中の一部を認識対象に指定した辞書

【図4】

## 大分類の方法

大分類方法1	マッチングの距離による大分類
大分類方法2	候補順位による大分類
大分類方法3	マッチングの距離と画像パラメータによる大分類
大分類方法4	組み合わせ大分類
大分類方法5	大分類を使わない

【図5】

辞書の種類	選択する大分類方法
小文字だけの辞書	大分類を行わない
英字、数字、仮名、漢字混在の辞書	マッチング距離、候補順位の両方による大分類



【図6】

